WHAT IS CLAIMED IS:

1. A system for generating a model, comprising:

a plurality of nodes, at least one of the nodes being configured to:

select a candidate condition,

request statistics associated with the candidate condition from other ones of the nodes,

receive the requested statistics from the other nodes,

form a rule based, at least in part, on the candidate condition and the requested statistics, and

selectively add the rule to the model.

2. The system of claim 1, further comprising:

a memory configured to store training data that includes a plurality of features; and

wherein the at least one node is configured to form the candidate condition from one or more of the features or complements of the features in the memory.

3. The system of claim 2, wherein the memory is further configured to store a plurality of instances, each of the instances corresponding to one or more of the features; and

wherein the at least one node is configured to:

identify a set of the instances that satisfy the candidate condition,

gather statistics regarding the set of instances, and

estimate a weight for the candidate condition based, at least in part, on the statistics.

4. The system of claim 3, wherein the at least one node includes a feature-to-instance index that maps the features to the instances in the memory that correspond to those features; and

wherein when identifying a set of the instances that satisfy the candidate condition, the at least one node is configured to use the feature-to-instance index.

5. The system of claim 3, wherein when gathering statistics regarding the set of instances, the at least one node is configured to determine one or more of first and second derivatives of a log likelihood for each of the instances in the set of instances.

6. The system of claim 2, wherein when requesting statistics associated with the candidate condition from other ones of the nodes, the at least one node is configured to:

generate a request that includes information associated with the candidate condition and the estimated weight, and

send the request to the other ones of the nodes.

7. The system of claim 6, wherein the other ones of the nodes are configured to:

generate the requested statistics based, at least in part, on the information associated with the candidate condition and the estimated weight, and

return the requested statistics to the at least one node.

8. The system of claim 7, wherein when generating the requested statistics, the other ones of the nodes are configured to determine one or more of first and second derivatives based,

at least in part, on the information associated with the candidate condition and the estimated weight.

9. The system of claim 1, wherein the at least one node is further configured to estimate a weight for the candidate condition prior to requesting statistics associated with the candidate condition.

10. The system of claim 9, wherein when requesting statistics associated with the candidate condition from other ones of the nodes, the at least one node is configured to:

generate a request that includes information associated with the candidate condition and the estimated weight, and

send the request to the other ones of the nodes.

11. The system of claim 10, wherein the other ones of the nodes are configured to:

generate the requested statistics based, at least in part, on the information associated with the candidate condition and the estimated weight, and

return the requested statistics to the at least one node.

12. The system of claim 1, wherein the other ones of the nodes are configured to:

generate the requested statistics based, at least in part, on information associated with the candidate condition, and

return the requested statistics to the at least one node.

13. The system of claim 12, wherein when generating the requested statistics, the other ones of the nodes are configured to generate one or more histograms associated with the candidate condition.

14. The system of claim 1, wherein the at least one node is further configured to determine a weight for the candidate condition based, at least in part, on the requested statistics.

15. The system of claim 14, wherein when forming a rule, the at least one node is configured to generate the rule based, at least in part, on the candidate condition and the weight for the candidate condition.

16. The system of claim 1, further comprising:

a memory configured to store training data; and

wherein when selectively adding the rule to the model, the at least one node is configured to add the rule to the model when the likelihood of the training data when the model includes the rule is sufficiently greater than when the model does not include the rule.

17. The system of claim 1, wherein the at least one node is further configured to transmit information regarding the rule to the other nodes.

18. The system of claim 1, wherein the at least one node is further configured to continue to select a candidate condition, request statistics, receive the requested statistics, form a rule, and selectively add the rule to the model for a number of iterations.

19. The system of claim 1, wherein the at least one node includes multiple ones of the nodes operating in parallel.

20. A method for generating a model, the method, performed substantially concurrently by a plurality of devices, comprising:

selecting candidate conditions;

requesting statistics associated with the candidate conditions from other ones of the devices;

receiving the requested statistics from the other devices;

forming rules based, at least in part, on the candidate conditions and the requested statistics; and

selectively adding the rules to the model.

21. A system for generating a model, comprising:

means for forming candidate conditions;

means for generating statistics associated with the candidate conditions;

means for forming rules based, at least in part, on the candidate conditions and the generated statistics; and

means for selectively adding the rules to the model.

22. A system for generating a model, comprising:

a repository configured to store training data that includes a plurality of features; and

a plurality of nodes configured to substantially concurrently:

select a candidate condition from one or more of the features or complements of the features,

request statistics associated with the candidate condition from other ones of the nodes,

receive the requested statistics from the other nodes,

form a rule based, at least in part, on the candidate condition and the requested statistics, and

selectively add the rule to the model based, at least in part, on a likelihood of the training data when the rule is added to the model.

23. The system of claim 22, wherein the repository includes a plurality of memory devices, each of the memory devices corresponding to one of the nodes.

24. A method for generating a model in a system that includes a plurality of nodes, comprising:

generating candidate conditions;

distributing the candidate conditions to the nodes;

generating statistics regarding the candidate conditions;

collecting the statistics for each of the candidate conditions at one of the nodes;

generating rules based, at least in part, on the statistics and the candidate conditions; and

selectively adding the rules to the model.

25. The method of claim 24, wherein the system further includes a memory that stores a plurality of instances; and

wherein the generating candidate conditions includes forming conditions that match at least a minimum number of the instances.

26. The method of claim 25, wherein the generating candidate conditions further includes:

determining a count value corresponding to a sum of a number of the instances that match the candidate conditions on all of the nodes, and

keeping the candidate conditions with corresponding count values greater than at least the minimum number.

27. The method of claim 25, wherein the generating candidate conditions further includes:

determining a count value corresponding to a number of the instances that match the candidate conditions on each of the nodes, and

keeping the candidate conditions with corresponding count values greater than at least the minimum number on one of the nodes.

28. The method of claim 24, wherein the system further includes a memory that stores a plurality of instances; and

wherein the generating statistics includes:

determining which of the candidate conditions match each of the instances,

forming condition-instance pairs based, at least in part, on a result of the determining,

sorting the condition-instance pairs to form a condition-instance list, and

generating statistics for each of the candidate conditions in the condition-instance list.

29. The method of claim 24, wherein the distributing the candidate conditions includes sending the candidate conditions to all of the nodes.

30. The method of claim 24, wherein the generating statistics includes generating one or more histograms for each of the candidate conditions.

31. The method of claim 24, wherein the collecting the statistics for each of the candidate conditions includes:

identifying the nodes to which to send the statistics for the candidate conditions, and

sending the statistics to the identified nodes.

32. The method of claim 31, wherein the identifying the nodes to which to send the statistics includes:

hashing the candidate conditions to form hash results, and

using the hash results to identify the nodes to which to send the corresponding candidate conditions.

33. The method of claim 24, further comprising:

estimating weights for the candidate conditions based, at least in part, on the statistics.

34. The method of claim 33, wherein the generating rules includes forming the rules based, at least in part, on the candidate conditions and the estimated weights.

35. The method of claim 24, wherein the system further includes a memory that stores training data; and

wherein the selectively adding the rules includes:

adding the rules to the model when a likelihood of the training data when the model includes the rules is sufficiently greater than when the model does not include the rules.

36. The system of claim 24, further comprising:

outputting the rules to other ones of the nodes when the rules are added to the model.

37. A system for generating a model, comprising:

means for generating new conditions;

means for distributing the new conditions to a plurality of nodes;

means for generating statistics regarding the new conditions at each of the nodes;

means for generating new rules based, at least in part, on the statistics and the new conditions; and

means for adding at least one of the new rules to the model.

38. A system for generating a model, comprising:

a plurality of nodes, at least one of the nodes being configured to:

receive candidate conditions,

generate statistics associated with at least a first one of the candidate conditions,

send the generated statistics to at least one other one of the nodes,

receive statistics regarding at least a second one of the candidate conditions from other ones of the nodes,

form a rule based, at least in part, on the second candidate condition and the received statistics, and

selectively add the rule to the model.